Migration Correlation: Definition and Efficient Estimation

P. Gagliardini[*], C. Gouriéroux[†]

First version: November 2003
This version: April 2004

[*]Università della Svizzera Italiana, and CREST. Corresponding author: Patrick Gagliardini, Institute of Finance, Università della Svizzera Italiana, Via Buffi 13, 6900 Lugano, Switzerland. E-mail: gagliarp@lu.unisi.ch. Tel: +41 91 912 4723. Fax: +41 91 912 4647.

[†]CREST, and University of Toronto.

We acknowledge G. Barone Adesi and two anonymous referees for very helpful comments.

Migration Correlation: Definition and Efficient Estimation

Abstract

The aim of this paper is to explain why cross-sectional estimated migration correlations displayed in the academic and professional literature can be either not consistent, or inefficient, and to discuss alternative approaches. The analysis relies on a model with stochastic migration in which the parameters of interest, that are migration correlations, are precisely defined. The impossibility of estimating consistently the migration correlations from cross-sectional data only is emphasized. We explain how to handle with individual rating histories, how to weight appropriately the cross-sectional estimators and how to estimate efficiently the joint migration probabilities at longer horizons.

Keywords: Credit Risk, Migration, Migration Correlation, Stochastic Transition, Rating.

JEL Number: C23, C35, G11

1 Introduction

The risk on a portfolio of corporate credits (including corporate bonds) depends on the level of individual default and recovery rates of the different borrowers, but also on their possible dependence. This dependence can be the consequence of structural links between the firms. For instance the failure of a firm can increase the default rate of its suppliers, which can imply chains of failures. This dependence can also be due to more general factors, which affect the general state of the economy, and the situation of all firms, or of a group of firms. For instance an increase of the exchange rate of the Euro versus \$ can diminish (resp. increase) the default rate of exporting (resp. importing) US firms.

Typically the total risk in a portfolio of credits (with positive allocations) is higher when the risks are positively correlated than when they can be assumed independent. Moreover this correlation effect cannot be suppressed by diversification. Even for portfolios including a large number of credits of similar types and sizes, the effect of the underlying common risk factors cannot be eliminated. Therefore the analysis of default correlation is a necessary step for understanding credit portfolio diversification and determining the capital required to compensate credit risk [see e.g. Gordy (1998), Nagpal, Bahar (1999), (2000), (2001)a, b, Lucas et alii (2001)]. In particular the importance of default correlation (and of migration correlation) has been recognized by the regulator and its value is a significant input in the computation of the required capital by means of the quantile of the portfolio credit losses (CreditVaR) [see The Basle Committee on Banking Supervision (2002)]. Since the CreditVaR is very sensitive to the value of the default correlation, this parameter has to be estimated carefully. An underestimation of the default correlation will imply an underestimation of the risk and a too low level of required capital. An unbiased, but inaccurate estimation of the default correlation will induce levels of required capital which are too erratic with respect to time.

Default correlations are generally deduced from data basis on default histories[1]. These basis correspond to panel data doubly indexed by individual (firm) and time. A special econometric literature is devoted to panel data. Indeed the total number of observations depend on the number of individuals (firms) in the panel [cross-sectional dimension] and of the number of obser-

[1] or rating histories for migration correlations.

vation dates [time dimension]. Thus a large number of observations can be obtained with a large cross-sectional dimension, a large time dimension, or both. The need for large cross-sectional or time dimension will depend on the parameters of interest. Loosely speaking a large cross-sectional dimension allows for accurate estimation of parameters with "individual" interpretation, but a large time dimension is also needed for the estimation of parameters with dynamic interpretations.

The aim of this paper is to discuss in the panel framework the standard descriptive estimation methods proposed in the literature to approximate default correlation or migration correlation [see e.g. Lucas (1995), Nagpal, Bahar (2001)b] and largely applied by practitioners and rating agencies for the computation of CreditVaR[2] [see e.g. de Servigny, Renault (2002)]. These methods are generally introduced without specifying a model driving transitions and, as we will see, they can be neither consistent, nor efficient. Our specific contribution is to provide a framework where a precise definition of migration correlation can be given and the consistency and efficiency of the estimators can be discussed. In addition, a more accurate estimator of migration correlations at longer horizons is presented.

The rest of the paper is organized as follows. In Section 2 the cross-sectional estimation method is reviewed and discussed. In particular its lack of consistency is pointed out. A precise definition of migration correlation is given in Section 3, by considering a model with stochastic transition matrix. The inconsistency of the cross-sectional estimator in this framework is discussed in Section 4. Then a consistent estimator of the migration correlation at horizon 1 is introduced in Section 5. In Section 6 it is explained how to deduce the migration correlations at longer horizons from the migration correlations at horizon 1. The finite sample properties of the different estimators are compared in a Monte-Carlo study presented in Section 7. Finally, Section 8 concludes.

[2]Other estimation approaches of default correlations are also considered in Finance. Some are based on the observation of equity prices, such as in the KMV approach, or on the observation of bond prices [Lando (1998), Duffie, Singleton (1999), Gourieroux, Monfort, Polimenis (2003)]. In particular they are used for pricing various credit derivatives proposed in the market. They can only be applied to large firms, quoted on the market, issuing bonds regularly. However for the regulator the CreditVaR has to be computed for the whole credit portfolio, including small or medium size firms, with a lot of over-the-counter credits. In this case the estimation approach from individual rating histories is the only possible one. The same remark applies for portfolios of retail credits including their derivatives, like the mortgage backed securities.

2 Calculating one year migration correlation [from Lucas (1995)]

In the financial literature the estimation of migration correlation has been introduced directly without defining precisely the parameter of interest and "without relying on a specific model driving transitions" [de Servigny, Renault (2002)] [3]. This section follows the usual presentation [see e.g. Lucas (1995), Nagpal, Bahar (2001)b]. The available data consists in a panel of rating histories:

$$Y_{i,t}, \quad i = 1, ..., n, \quad t = 1, ..., T,$$

where i denotes the individual (either the firm, or the bond) and t the year. The variable $Y_{i,t}$ is polytomous qualitative, with alternatives $k = 1, ..., K$, indicating the different admissible grades. In general one of these alternatives corresponds to default, $k = K$ say. For instance there are eight grades for Standard & Poor's from the highest rating AAA to debt in default payment D, when the ratings C, CC, CCC are aggregated. The cross-sectional dimension is from $n = 10000$ for Moody's, S&P's where the data concern large international firms, up to $n = 130000$ for the French Central Bank, which follows all French firms including small and medium size firms. The time unit is generally the year, leading to a time dimension of $T = 15 - 20$. Thus the cross-sectional dimension is much larger than the time dimension [see Foulcher, Gouriéroux, Tiomo (2003) for a comparison of the ratings of the main rating agencies Moody's, S & P's, Fitch and of the French Central Bank].

At any given year t, we can compute the structure of individuals (firms or bonds) per rating grade: $(N_{k,t}, k = 1, ..., K)$, where $N_{k,t}$ denotes the number of firms (bonds) in grade k at t. We can also consider the transitions. If $N_{kl,t},\ k, l = 1, ..., K$, denote the number of firms (bonds) migrating from rating class k at $t-1$ to rating class l at date t, the transition probabilities are approximated by:

$$\widehat{\pi}_{kl,t} = \frac{N_{kl,t}}{N_{k,t-1}}, \quad \text{(say)}.$$

Such matrices are regularly reported by Moody's, S & P, the French Central Bank, etc [see e.g. Brady, Bos (2002), Brady, Vazza, Bos (2003), Bardos,

[3] This can explain the following remark by Lucas (1995) p82: "These historical statistics describe only observed phenomena, not the true underlying correlation relationship".

Foulcher, Bataille (2004)]. They summarize the individual migrations of a given firm. Such an estimated transition matrix is given in Table 1.

It is usually considered that joint migrations can be analyzed in a similar way, by calculating the ratio between the number of pairs of firms (bonds) in a given rating class, which actually migrated to a given category (for example default), and the total number of pairs of firms (bonds) in the rating class. Then an estimator is defined by:

$$\widehat{p}_{kl,t} = \frac{N_{kl,t}\left(N_{kl,t}-1\right)}{N_{k,t-1}\left(N_{k,t-1}-1\right)}, \tag{1}$$

where k [resp. l] is the starting grade [resp. the final grade]. The intuitive idea is to approximate a "probability of joint migration" by drawing the pairs of bonds without replacement in the population and taking the empirical counterpart. This type of estimator is standard in survey sampling theory, where it is used to estimate the so-called second order inclusion probability involved in the Horvitz-Thompson estimator [see e.g. Kish (1967), Konijn (1973)]. However it is clearly inappropriate in the framework of default correlation. Let us study what the quantity $\widehat{p}_{kl,t}$ is really approximating. It is immediately realized that, for a large portfolio ($n \sim \infty$), both $N_{k,t-1}$ and $N_{kl,t}$ will be large and:

$$\widehat{p}_{kl,t} \;\sim\; \widehat{\pi}^2_{kl,t}.$$

Thus, in our framework of large cross-sectional dimension, the estimators $\widehat{p}_{kl,t},\ k,l = 1,...,K$, bring no additional information than the estimators $\widehat{\pi}_{kl,t},\ k,l = 1,...,K$, of the individual transitions and in particular cannot be used to measure a notion of joint migration. A similar argument can be given for the migration correlations. The cross-sectional estimator is given by:

$$\widehat{\rho}_{kl,t} = \frac{\widehat{p}_{kl,t} - \left(\widehat{\pi}_{kl,t}\right)^2}{\widehat{\pi}_{kl,t}\left(1-\widehat{\pi}_{kl,t}\right)}. \tag{2}$$

If the size n of the portfolio is large, estimator $\widehat{p}_{kl,t}$ is close to $\left(\widehat{\pi}_{kl,t}\right)^2$, and $\widehat{\rho}_{kl,t}$ is close to zero, no matter of the true value of the underlying migration correlation[4].

This argument[5] will be developed in greater details in the next section,

[4]Estimator $\widehat{\rho}_{kl,t}$ is numerically equal to $-1/(N_{k,t-1}-1)$.

[5]The same argument holds for the modified estimator of de Servigny, Renault (2002) corresponding to drawing with replacement, introduced to avoid negative estimated probabilities of joint migration.

where a model with stochastic transition is introduced and used to define the parameters of interest, that are the probabilities of joint migration or the migration correlations.

3 Stochastic migration intensity

A basic model for understanding migration dependence is the extension of the stochastic intensity model introduced for default risk [see e.g. Lando (1998), Duffie, Singleton (1999), Gouriéroux, Monfort, Polimenis (2003)]. This specification assumes that the rating chains $(Y_{i,t},\ t \text{ varying})$, $i = 1, ..., n$, are independent identically distributed Markov processes, when the common transition matrices (Π_t), with elements $\pi_{kl,t}$, are given, and the model is completed by assuming that the transition matrices are stochastic. For expository purpose, we assume[6]:

Assumption A.1: The transition matrices at horizon 1 are stochastic, independent with identical distributions (i.i.d.).

Thus the joint (serial and instantaneous) dependence between ratings passes through the stochastic transition. This explains why it is not possible to identify "whether historical fluctuations in default rates are caused by default correlation or simply by (stochastic) changes in default probabilities" [Lucas (1995), p82]. We get a nonlinear factor model in which the factors are the elements of the transition matrices. The number of independent factors is $(K-1)^2$ due to the unit mass restrictions on transition probabilities and the interpretation of default as an absorbing barrier.

Since the underlying time dependent transition matrices are not observable, the joint distribution of rating chains has to be derived by integrating out the unobservable factor Π_t. When the transitions have been integrated out, the joint rating vector $(Y_{1,t}, ..., Y_{n,t})'$ still defines an homogeneous Markov chain [see Appendix 1] with K^n admissible states. In practice $K = 10$, the number of firms is several thousands (about 10000 for the data base of S & P, 130000 for the data base of the French Central Bank) and the state space

[6] See Gagliardini, Gouriéroux (2003) for the general framework of serially dependent stochastic transition matrices. For serially dependent transition matrices, the migration correlation has to be defined conditional on the whole rating histories, not only the last rating; but consistent estimators of the path dependent migration correlation can also be derived, and differ from the estimator introduced in this paper.

dimension of the joint rating chain is for instance 10^{10000} [7]. Therefore the transitions corresponding to the joint rating processes are difficult to display, even in the framework of Assumption A.1. However this framework is appropriate for a first analysis of joint migration of a pair of firms (or bonds). For instance there is dependence between the ratings of two different firms (bonds), given their lagged ratings. The migration dependence is a consequence of the unobservable factor which jointly influences the ratings of all firms (bonds). Typically the realization at date t of the transition matrix Π_t can correspond to a matrix with larger probabilities of downgrades (resp. upgrades), which will imply more (joint) downgrades (resp. upgrades) of individual ratings. This effect is illustrated below in Figure 1, where two rating histories are displayed[8].

[Figure 1: Simulated rating histories]

Due to a regime switching of the transition matrix, we observe some underlying effect towards downgrades at dates 7 and 13. This implies joint downgrades, but with different sensitivities for the two firms.

In the stochastic transition framework it is now possible to define precisely the parameters of interest. For this purpose let us focus on a pair of firms (bonds) i, j [9]. The bivariate process $(Y_{i,t}, Y_{j,t})$ is also an homogeneous Markov process [see Appendix 1] with state space including the admissible pairs of grades (k, l). Let us consider its transition matrix. The elements of this matrix are joint transition probabilities defined by:

$$p_{kk^*,ll^*} = P\left[Y_{i,t+1} = k^*, Y_{j,t+1} = l^* \mid Y_{i,t} = k, Y_{j,t} = l\right]. \tag{3}$$

These joint transition probabilities do not depend on the selected couple of firms (bonds) (i, j). Moreover they admit a simple expression in terms of the stochastic transition probabilities[10]:

$$p_{kk^*,ll^*} = E\left[\pi_{kk^*,t}\pi_{ll^*,t}\right]. \tag{4}$$

[7] The number of firms is still large if the homogeneity (i.i.d.) assumption is introduced on subsets of firms to account for region or industry specific effects [see e.g. Nagpal, Bahar (2001)].

[8] The simulation is based on the three states ordered probit model presented in Section 7.

[9] It would be also possible to consider joint migration for three, four, etc, firms in order to study migration clustering phenomenon [see Gouriéroux, Monfort (2002) for the analysis of clustering for default risk].

[10] The expressions of the joint transition probabilities are modified when the stochastic

Instead of the joint migration transition matrix, it is also possible to define the migration correlation[11]:

$$\rho_{kk^*,ll^*} = Corr\left[\mathbb{I}_{Y_{i,t+1}=k^*}, \mathbb{I}_{Y_{j,t+1}=l^*} \mid Y_{i,t}=k, Y_{j,t}=l\right], \tag{5}$$

where $\mathbb{I}_{Y=k} = 1$, if $Y = k$, and 0 otherwise. The migration correlation is equal to:

$$\rho_{kk^*,ll^*} = \frac{p_{kk^*,ll^*} - \alpha_{kk^*}\alpha_{ll^*}}{\sqrt{\alpha_{kk^*}\left(1-\alpha_{kk^*}\right)}\sqrt{\alpha_{ll^*}\left(1-\alpha_{ll^*}\right)}}, \tag{6}$$

where: $\alpha_{kk^*} = E\left[\pi_{kk^*,t}\right]$ is the expected transition probability. The migration correlation can be written as:

$$\begin{aligned}\rho_{kk^*,ll^*} &= \frac{cov\left(\pi_{kk^*,t}, \pi_{ll^*,t}\right)}{\sqrt{\alpha_{kk^*}\left(1-\alpha_{kk^*}\right)}\sqrt{\alpha_{ll^*}\left(1-\alpha_{ll^*}\right)}} \\ &= corr\left(\pi_{kk^*,t}, \pi_{ll^*,t}\right)\sqrt{\frac{V\left(\pi_{kk^*,t}\right)}{\alpha_{kk^*}\left(1-\alpha_{kk^*}\right)}}\sqrt{\frac{V\left(\pi_{ll^*,t}\right)}{\alpha_{ll^*}\left(1-\alpha_{ll^*}\right)}},\end{aligned}$$

where $V(\pi)$ denotes the variance of π. In the latter decomposition, the first component $corr\left(\pi_{kk^*,t}, \pi_{ll^*,t}\right)$ measures the link between the underlying transition probabilities. The two other terms account for the relationship between the discrete rating indicators $\mathbb{I}_{Y=k}$ and the continuous transition probabilities.

To summarize, joint (bivariate) migrations can be equivalently described in terms of the (K^2, K^2) matrices p or ρ. All joint migrations provide information on the risk of a credit portfolio, and have to be estimated. However for illustration some specific ones are often presented and discussed. For instance the joint migration towards default corresponds to $k^* = l^* = K$, and is summarized by K^2 different numbers. The up-up migrations corresponds to $k^* = k+1$, $l^* = l+1$, that is to joint upgrade by one tick. There exist $(K-1)^2$ such measures. Similarly it is possible to define down-down, and up-down migration probabilities. At this step it is important to note that the

transition matrices are serially correlated. Indeed the expectation has to be evaluated conditional on the whole individual rating histories. Since the parameter of interest has a different expression, its consistent estimator has to be modified accordingly [see Gagliardini, Gouriéroux (2003)].

[11] The joint transition probabilities p_{kk^*,ll^*} are often called migration correlations in the literature. It seems important to distinguish the definitions of p and ρ to avoid misleading interpretations.

Basle Committee has recently proposed to account for migration correlation, but has only proposed a model when $k^* = l^*$, $k = l$, which is not sufficient to take account of all dependences [see The Basle Committee on Banking Supervision (2002)].

At the end of Section 3 and under Assumption A.1, the parameters of interest either the joint migration probabilities p_{kk^*,ll^*}, or the associated migration correlations ρ_{kk^*,ll^*} are clearly defined.

4 Non-consistency of the cross-sectional estimator

Let us explain the reason for the inconsistency of the cross-sectional estimator, when the cross-sectional dimension n is large. As mentioned in Section 2, for large portfolios ($n \sim \infty$), the cross-sectional estimator $\widehat{p}_{kl,t}$ converges to $\pi^2_{kl,t}$ and not to the parameter of interest $p_{kl,kl}$ corresponding to the diagonal term defined in equation (3). This result is a consequence of the factor representation of the rating processes. Conditional on $\Pi_t, \Pi_{t-1}, ...$ the individual rating chains are independent identically distributed. Thus the strong law of large numbers can be applied conditional on the factors and we get:

$$\lim_{n\to\infty} \widehat{p}_{kl,t} = \pi^2_{kl,t}. \tag{7}$$

The possibility of applying the law of large numbers and the central limit theorem conditional on the factors corresponds to the so-called infinitely fine grained assumption proposed by the Basle Committee. In particular, if the number of firms (bonds) n is large, the realization of the transition matrix at date t will be completely known. In fact it is easily seen that the set of empirical transitions $(\widehat{\pi}_{kl,t},\ k, l = 1, ..., K,\ t = 1, ..., T)$ is a sufficient statistics for the model with stochastic transition. From the cross-sectional information at year t, it is at best possible to reconstitute the transition matrix Π_t. But it is not possible to reconstitute the matrices p and ρ. In financial terms the systematic risk Π_t cannot be eliminated by a cross-sectional averaging, that is by individual diversification. For instance an unbiased estimation of a covariance $Cov\,(\pi_{kk^*,t}, \pi_{ll^*,t})$ requires at least two dated observations of the transition, that is three successive years of data, and a pure cross-sectional approach cannot allow for identifying joint migrations.

At a first sight it can be surprising that the cross-obligor default correlation cannot be estimated consistently from an infinite number n of (cross-sectional) data. In fact it is important to note that different asymptotic theories can be considered in a panel framework according to the dimension, either n, or T, which tends to infinity. Some parameters, such as the transition matrix of date t, can be estimated consistently when n tends to infinity and only one observation date (that is t) is available. Other parameters, such as the migration correlations, have an interpretation involving time and need also T tending to infinity to be consistent[12].

The same argument applies to the maximum likelihood estimator considered by Gordy, Heitfield (2002) in a model where $\Pi_t = \Pi$ is time independent and stochastic. The time dimension is not sufficient to get consistency since the data are both with equal correlation with respect to time and individual.

5 A time averaged estimator

Let us assume a rather large portfolio ($n/K \geq 30$ approximately) and several years of observations. In a first step the underlying transition probabilities $(\pi_{kl,t})$ will be replaced by their cross-sectional empirical counterparts $(\widehat{\pi}_{kl,t})$. Then the parameter of interest p or ρ will be approximated by averaging on time. More precisely, from equation (4), the joint migration probability p_{kk^*,ll^*} is the common mean of all the products $\pi_{kk^*,t}\pi_{ll^*,t}$, $t = 1, ..., T$. For a large portfolio, we have [by central limit theorem applied conditional on (Π_t)]:

$$\widehat{\pi}_{kk^*,t}\widehat{\pi}_{ll^*,t} \simeq \pi_{kk^*,t}\pi_{ll^*,t} + u_{kk^*,ll^*,t}, \quad t = 2, ..., T, \tag{8}$$

where $V(u_{kk^*,ll^*,t}) = \omega_{kk^*,ll^*}$, say, is independent of the year. We deduce that:

$$\widehat{\pi}_{kk^*,t}\widehat{\pi}_{ll^*,t} \simeq p_{kk^*,ll^*} + u_{kk^*,ll^*,t} + v_{kk^*,ll^*,t}, \quad t = 2, ..., T, \tag{9}$$

where $v_{kk^*,ll^*,t} = \pi_{kk^*,t}\pi_{ll^*,t} - p_{kk^*,ll^*}$ are i.i.d. errors with zero-mean and a variance independent of t (Assumption A.1). Moreover the errors u and v are independent. The relation (9) defines a linear model, with the constant

[12] Generally the consistency of an estimator requires some geometric ergodicity condition on the observations, that is a correlation between observations tending to zero at a geometric decay rate with the distance between the observation indexes: $\exists r, 0 \leq r < 1$, $c > 0$ such that $|corr(Y_{i,t}, Y_{j,t}) \mid Y_{i,t-1} = k, Y_{j,t-1} = l| \leq cr^{j-i}$. This condition is typically not satisfied in the present framework where the cross-sectional correlation is constant.

as explanatory variable, p_{kk^*,ll^*} as the parameter. Since the model is homoscedastic, the optimal weights to compute the least squares estimator are identical. Thus the estimator is:

$$\widehat{p}_{kk^*,ll^*} = \frac{1}{T-1} \sum_{t=2}^{T} \widehat{\pi}_{kk^*,t} \widehat{\pi}_{ll^*,t}. \tag{10}$$

The estimator of the migration correlation is [see equation (6)]:

$$\widehat{\rho}_{kk^*,ll^*} = \frac{\widehat{p}_{kk^*,ll^*} - \widehat{\alpha}_{kk^*} \widehat{\alpha}_{ll^*}}{\sqrt{\widehat{\alpha}_{kk^*} \left(1 - \widehat{\alpha}_{kk^*}\right)} \sqrt{\widehat{\alpha}_{ll^*} \left(1 - \widehat{\alpha}_{ll^*}\right)}}, \tag{11}$$

where $\widehat{\alpha}_{kk^*} = \frac{1}{T-1} \sum_{t=2}^{T} \widehat{\pi}_{kk^*,t}$. The estimators of matrices p or ρ are consistent when both the cross sectional and time dimensions, n and T, respectively, tend to infinity. When n is large $\widehat{\pi}_{kk^*,t} \widehat{\pi}_{ll^*,t}$ is close to $\pi_{kk^*,t} \pi_{ll^*,t}$ and the approximated linear model (9) is valid. For n large, the OLS estimator is unbiased, efficient (among linear estimators). However its variance can be far from zero if T is small. Thus the additional condition $T \to \infty$ is needed to get the consistency of the OLS with T aggregated transitions.

The consistency result is valid if the Markov chain is recurrent ergodic. It is important to note that the recurrence condition, that is the fact that the chain visits each state an infinite number of dates, is not satisfied when the state space includes an absorbing barrier, such as default. However the recurrence condition can be recovered if the population of interest is renewed at each date. More precisely it has to be assumed that some new firms are created to balance the defaulted individuals. This latter assumption is approximately satisfied for the data base of the rating agencies and of the French Central Bank [see Gagliardini, Gouriéroux (2003) for a more detailed discussion].

Estimators of the joint (bivariate) transition probability averaged on several dates have already been considered in the literature [see e.g. Bahar, Nagpal (2001)b, de Servigny, Renault (2002)], without underlying statistical model and without any justification in terms of consistency or efficiency. They have often been defined with different weights for the different dates. For instance the estimator of $p_{kl,kl}$ introduced by Bahar, Nagpal (2001)b and by de Servigny, Renault (2002) are considered below. The estimator by Bahar, Nagpal (2001)b is:

$$\widetilde{p}_{kl,kl} = \frac{\sum_{t=2}^{T} N_{kl,t} \left(N_{kl,t} - 1\right)}{\sum_{t=2}^{T} N_{k,t-1} \left(N_{k,t-1} - 1\right)} \sim \frac{\sum_{t=2}^{T} N_{kl,t}^2}{\sum_{t=2}^{T} N_{k,t-1}^2} = \sum_{t=2}^{T} \frac{N_{k,t-1}^2}{\sum_{t=2}^{T} N_{k,t-1}^2} \widehat{\pi}_{kl,t}^2.$$

It is easily checked that, if n is large, and if the population is renewed to ensure the recurrence of the chain, we get:

$$N_{k,t-1}^2 \sim n^2 p_k^2, \quad \sum_{t=2}^{T} N_{k,t-1}^2 \sim (T-1)\, n^2 p_k^2,$$

where (p_k) is the marginal distribution of the chain, and therefore $\widetilde{p}_{kl,kl} \sim \widehat{p}_{kl,kl}$. Thus the introduction of weights is absolutely not necessary. Similarly the estimator proposed by de Servigny, Renault (2002) is given by:

$$\sum_{t=2}^{T} \frac{N_{k,t-1}}{\sum_{t=2}^{T} N_{k,t-1}} \widehat{\pi}_{kl,t}^2,$$

and the same remark applies.

6 Longer horizons

It is a common practice to use the estimation formula (10) simultaneously for different horizons [see e.g. Nagpal, Bahar (2001)]. For instance the joint bivariate transition probability at horizon 2:

$$p_{kk^*,ll^*}(2) = P\left[Y_{i,t+2} = k^*, Y_{j,t+2} = l^* \mid Y_{i,t} = k, Y_{i,t} = l\right], \tag{12}$$

is estimated by:

$$\widetilde{p}_{kk^*,ll^*}(2) = \frac{1}{T-2} \sum_{t=3}^{T} \widehat{\pi}_{kk^*,t}(2) \widehat{\pi}_{ll^*,t}(2), \tag{13}$$

where $\widehat{\pi}_{kk^*,t}(2)$ denotes the empirical cross-sectional transition probability for period $(t-2, t)$. It is important to note that it is not appropriate to use simultaneously the empirical estimator of the joint transition probability for different horizons. Indeed, whereas $\widehat{p}_{kk^*,ll^*}$ is efficient for p_{kk^*,ll^*}, the estimator $\widetilde{p}_{kk^*,ll^*}(2)$ can be improved, by taking into account the information included in the statistical model. Indeed, under Assumption A.1, the transition matrices at horizon 2 such as $\Pi_t(2) = \Pi_{t-1}\Pi_t$ and $\Pi_{t+1}(2) = \Pi_t\Pi_{t+1}$ depend on the common matrix Π_t and therefore are dependent. Thus Assumption A.1 cannot be satisfied simultaneously at horizons 1 and 2. Under Assumption A.1 the estimators $\widehat{\Pi}_t(2)$ and $\widehat{\Pi}_{t+1}(2)$ are also dependent, since they are computed on overlapping periods. Loosely speaking the estimator $\widetilde{p}$ computed from $T = 20$ observations for horizon 7 [see e.g. Nagpal, Bahar (2001)] has an accuracy equivalent to an estimator under independence computed from about $20/7 \sim 3$ observations. Therefore it is not very accurate due to the non-consistency of cross-sectional estimators [13].

A better accuracy can be recovered along the following lines. By the law of iterated expectations the joint transition probability at horizon 2 can be written as:

$$\begin{aligned} p_{kk^*,ll^*}(2) &= E\left[\pi_{kk^*,t}(2)\pi_{ll^*,t}(2)\right] \\ &= \sum_m \sum_n E\left[\pi_{km,t-1}\pi_{mk^*,t}\pi_{ln,t-1}\pi_{nl^*,t}\right] \\ &= \sum_m \sum_n E\left[\pi_{km,t-1}\pi_{ln,t-1}\right] E\left[\pi_{mk^*,t}\pi_{nl^*,t}\right] \\ &= \sum_m \sum_n p_{km,ln} p_{mk^*,nl^*}. \end{aligned}$$

An estimator with an accuracy of the same order of magnitude as $\widehat{p}_{kk^*,ll^*}$, that is taking into account the whole time dimension, and which is asymptotically efficient is:

$$\widehat{p}_{kk^*,ll^*}(2) = \sum_m \sum_n \widehat{p}_{km,ln}\widehat{p}_{mk^*,nl^*}, \tag{14}$$

where $\widehat{p}_{km,ln}$ is given by formula (10) applied to horizon 1. This approach is easily understood if it is noted that under Assumption A.1 the bivariate process $(Y_{i,t}, Y_{j,t})$ is a Markov process of order 1 with transition matrix:

$$P = \left[p_{kk^*,ll^*}\right].$$

Therefore its transition matrix at horizon 2 is:

$$P(2) = P^2,$$

which is precisely the formula used in (14).

Estimators for any horizon h corresponding to (13) and (14) are derived in a similar way by formula $P(h) = P^h$.

[13]The same remark applies to the ten-year default correlations estimated in Lucas (1995).

Finally, estimators (13) and (14), and their generalizations for longer horizons, can be used to define estimators for migration correlations at horizon h:

$$\widetilde{\rho}_{kk^*,ll^*}(h) = \frac{\widetilde{p}_{kk^*,ll^*}(h) - \widetilde{\alpha}_{kk^*}(h)\widetilde{\alpha}_{ll^*}(h)}{\sqrt{\widetilde{\alpha}_{kk^*}(h)\left[1-\widetilde{\alpha}_{kk^*}(h)\right]}\sqrt{\widetilde{\alpha}_{ll^*}(h)\left[1-\widetilde{\alpha}_{ll^*}(h)\right]}}, \qquad (15)$$

where $\widetilde{\alpha}_{kk^*}(h) = \frac{1}{T-h}\sum_{t=h+1}^{T}\widehat{\pi}_{kk^*,t}(h)$, and:

$$\widehat{\rho}_{kk^*,ll^*}(h) = \frac{\widehat{p}_{kk^*,ll^*}(h) - \widehat{\alpha}_{kk^*}(h)\widehat{\alpha}_{ll^*}(h)}{\sqrt{\widehat{\alpha}_{kk^*}(h)\left[1-\widehat{\alpha}_{kk^*}(h)\right]}\sqrt{\widehat{\alpha}_{ll^*}(h)\left[1-\widehat{\alpha}_{ll^*}(h)\right]}}, \qquad (16)$$

where $\widehat{\alpha}(h) = \widehat{\alpha}^h$.

It has been usual to compare the migration correlations estimated for different horizons by the crude empirical estimation method. It is clear that such a comparison has to be performed carefully since the migration correlations do not have the same interpretations for different terms. The formula $P(h) = P^h$ derived under Assumption A.1 explains how to correct the estimated probabilities of joint transitions for the term effect. Loosely speaking the matrices to be compared are not $\widetilde{P}(h)$, but rather $\widetilde{P}(h)^{1/h}$, in order to check for a "flat term structure" of migration correlation.

7 An illustration

In order to highlight the difference between estimation approaches at the different horizons, a Monte-Carlo study is performed in this section. Let us first present the model for the stochastic transition matrices (Π_t) used for the simulation, and then discuss the Monte-Carlo results.

7.1 An ordered probit model for the stochastic transition matrices

A standard specification for the transition matrices proposed in the academic literature [see e.g. Bangia et alii (2002), Albanese et alii (2003), Feng, Gouriéroux, Jasiak (2003)] and often adopted by market practice [see e.g. Gupton, Finger, Bhatia (1997), Crouhy, Galai, Mark (2000)] is the ordered probit model. This approach assumes a continuous quantitative grade for each firm, which is used to define the qualitative ratings. Let us denote by s_{it} the continuous latent grade of firm (bond) i at date t. Such a latent grade is sometimes computed regularly by the rating specialists, especially for internal ratings. Generally it is confidential and has to be considered as unobservable. In other approaches based on Merton's model [see Merton (1974)], this latent grade is defined as the ratio of asset value and liabilities[14]. It is also unobservable, whenever a detailed balance sheet of the firm is unknown. The ordered probit model [see e.g. Maddala (1986), Gouriéroux (2000)] is defined by i) specifying the latent model, that is the dynamics of the latent variables, and ii) by explaining how the observable endogenous variables, that are the qualitative ratings, are related to the underlying continuous grades.

i) The latent model

Let us assume known the ratings of the different firms (bonds) at date $t-1$. Given these ratings, we assume that the underlying grades s_{it}, $i = 1, ..., n$, $t = 1, ..., T$, can be written as:

$$s_{it} = \alpha_k + Z_t + \varepsilon_{it},$$

where the errors (ε_{it}), $i = 1, ..., n$, and the common factor (Z_t) are independent standard gaussian variables, and the intercept α_k depends on the lagged rating $Y_{i,t-1} = k$ of firm (bond) i. Note that the conditional distribution of the score of a firm given the past depends on the previous rating of the firm only.

ii) The link between the latent and observed variables

Let us now consider a model with $K = 3$ states, where $k = 3$ denotes the default absorbing barrier. The qualitative ratings at date t are deduced by discretizing the underlying continuous grades:

$$Y_{it} = l, \quad \text{iff} \quad c_{l-1} \leq s_{it} < c_l,$$

where $c_0 = -\infty < c_1 < c_2 < c_3 = \infty$ are given thresholds. Then the transition matrix is given by:

$$\Pi_t = \begin{pmatrix} \Phi(a_{11} - Z_t) & \Phi(a_{12} - Z_t) - \Phi(a_{11} - Z_t) & 1 - \Phi(a_{12} - Z_t) \\ \Phi(a_{21} - Z_t) & \Phi(a_{22} - Z_t) - \Phi(a_{21} - Z_t) & 1 - \Phi(a_{22} - Z_t) \\ 0 & 0 & 1 \end{pmatrix},$$

[14] In other words Merton's model corresponds to a crude specification of the score, where the score coincides with this ratio. In the usual scoring methodology this ratio is one explanatory variable included among several other ones.

where the parameters $a_{kl} = c_l - a_k$, $k, l = 1, 2$, are set to[15]:

$$a_{11} = 1, a_{12} = 4, a_{21} = -1, a_{22} = 2 \ .$$

Thus the first two rows of the transition matrix correspond to an ordered probit model, with a common latent variable. This type of probit model has been suggested for the analysis of credit risk. It has been proposed to identify the common factor with one macroeconomic variable as a cycle indicator [see e.g. Bangia et alii (2003)] [16]. This interpretation is difficult to test from the available data which concern a limited period of time. Moreover the model has to be completed by a dynamic model for cycles. When the factor is let unspecified as in the present model, the transition matrix becomes stochastic by means of Z_t. The distribution of the transition matrix Π_t, that is the distribution of the four independent transition probabilities, has a rather complicated form. However it is easy to derive its first and second moments by simulation. For this purpose the number of replications has been fixed to 1000000. The expected transition matrix $\alpha = E\left(\Pi_t\right)$ is given by:

$$\alpha = \begin{pmatrix} 0.761 & 0.237 & 0.002 \\ 0.240 & 0.682 & 0.078 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, as observed in practice, the diagonal elements are rather large. Moreover the values 0.237 and 0.240, corresponding to one tick downgrade and upgrade, are of the same order of magnitude. The class $k = 1$ is less risky than $k = 2$ as revealed by the ranking of the associated default probabilities. Let us now consider the variance-covariance matrices of the first and second rows, $\pi_{1,t}$ and $\pi_{2,t}$, of Π_t. They are are given by:

$$V\left(\pi_{1,t}\right) = \begin{pmatrix} 0.056 & -0.054 & -0.002 \\ -0.054 & 0.053 & 0.001 \\ -0.002 & 0.001 & 0.001 \end{pmatrix}, \quad V\left(\pi_{2,t}\right) = \begin{pmatrix} 0.055 & -0.039 & -0.016 \\ -0.039 & 0.040 & -0.001 \\ -0.016 & -0.001 & 0.017 \end{pmatrix},$$

whereas their covariance is:

$$Cov\left(\pi_{1,t}, \pi_{2,t}\right) = \begin{pmatrix} 0.042 & -0.014 & -0.028 \\ -0.041 & 0.014 & 0.027 \\ -0.001 & -0.000 & 0.001 \end{pmatrix}.$$

[15] The parameters satisfy the constraint: $a_{11} + a_{22} = a_{12} + a_{21}$.

[16] Under this interpretation, it would be necessary to assume serially dependent transition matrices to get dynamics of switching regimes.

Since the rows of the transition matrix sum up to 1, the rows and columns of the variance and covariance matrices sum up to zero. Since the diagonal elements of the variance-covariance matrices have to be positive, at least one correlation is negative. Thus the unit mass restriction can be a source of misleading interpretations of the migration correlation sign.

7.2 Results

A Monte-Carlo study has been performed to analyze the finite sample properties of the joint migration probability and migration correlation estimators discussed in Sections 2, 5 and 6. It is based on $S = 10000$ replications of the rating transitions of $n = 1000$ firms over $T = 20$ years in the ordered probit model defined in Section 7.1. Note that the number of firms is kept fixed. The population is not renewed by introducing new created firms to balance default. Even if the estimators are not consistent without renewing the population, it is still possible to study the finite sample properties of the estimators.

7.2.1 Migration correlations at horizon 1

Table 2 displays the estimated joint (bivariate) transition probabilities at horizon 1 obtained by using the time averaged estimator. We focus on the case $k = l$, $k^* = l^*$, which corresponds to transitions with identical starting rating classes and identical final rating classes for the pair of firms. In Table 3 we provide the corresponding results obtained using the cross-sectional estimator evaluated at date $t = 10$. For both estimators we report the average, the median, the standard deviation, the Mean Squared Error (MSE), and the range between the α and $1 - \alpha$ quantiles, for $\alpha = 5\%$ and 1%. The true value of the corresponding parameters of interest are reported in Table 4.

For both estimators the expectation is very close to the true parameter value, pointing out that they are both unbiased for the selected sample sizes. However, the two estimators strongly differ in terms of accuracy: the time averaged estimator features better accuracy with respect to the pure cross-sectional one. This is immediately seen by comparing the standard deviations. For the estimator proposed in this paper, they are smaller by a factor of about $\sqrt{20} \sim 4.5$. These findings are reinforced by comparing the interquantile ranges, for $\alpha = 0.05, 0.01$. These intervals are much tighter for the time averaged estimator, indicating that its distribution is more concentrated around the mean value. In order to better visualize these results, Figures 2 and 3 display the histograms of the time averaged and cross-sectional estimators, respectively, for the following transitions: $(1,1) \rightarrow (2,2)$ (joint downgrade), $(1,1) \rightarrow (3,3)$ (joint default starting from the better rating), $(2,2) \rightarrow (1,1)$ (joint upgrade), and $(2,2) \rightarrow (3,3)$ (joint default starting from the worst rating).

[Insert Figure 2: Histogram of the distribution of the time averaged estimator]

[Insert Figure 3: Histogram of the distribution of the cross-sectional estimator]

The distributions corresponding to the cross-sectional estimator are much more dispersed and feature fatter tails. In particular, the distributions of the estimator admit a peak at 0, which implies a significant probability of strongly underestimating the joint migrations. The same type of feature will of course be observed with a time averaged estimator when the number of observation dates is too small.

Similar tables can be derived for the time averaged and cross-sectional estimators of the migration correlation matrix ρ. The estimates by the time averaged method (resp. the cross-sectional estimator) are provided in Table 5 (resp. 6), whereas the true correlations are given in Table 7. It is immediately seen that the cross-sectional estimator of the ρ matrix is highly biased, with values uniformly close to zero.

To summarize, the Monte-Carlo study shows that the theoretical inconsistency of the cross-sectional estimator has serious consequences for any time averaged estimator with small finite time dimension T (even if the population is not renewed), such as very erratic estimates and strong underestimation of the probability of joint migrations. This bias has direct consequences for risk management. Indeed by underestimating the joint migrations we underestimate the default correlation and thus the risk included in the credit portfolio.

7.2.2 Longer horizons

Let us now consider the estimation of migration correlations at horizon longer than 1. Table 8 displays the estimated bivariate transition probabilities at horizon $h = 7$ obtained using the standard estimator $\widetilde{p}_{kl,kl}(7)$ defined in equation (13). The corresponding results for estimator $\widehat{p}_{kl,kl}(7)$ taking into account the Markov property [see equation (14)] are displayed in Table 9. The true value of the parameters of interest are reported in Table 10.

Let us first consider the mean of the estimators. Both estimators feature a moderate finite sample bias, which is in general smaller for estimator $\widetilde{p}_{kl,kl}(7)$ [17]. However, estimator $\widehat{p}_{kl,kl}(7)$ is preferable in terms of the median. Indeed, for both estimators the median is often below the mean, but the median of $\widehat{p}_{kl,kl}(7)$ is generally closer to the true parameter value. In particular, the median of $\widetilde{p}_{kl,kl}(7)$ often underestimates quite severely the true joint migration probability, especially in the cases of joint default. These features are confirmed by the empirical distributions of the two estimators, which are reported in Figures 4 and 5.

[Insert Figure 4: Histogram of the distribution of $\widetilde{p}_{kl,kl}(7)$]

[Insert Figure 5: Histogram of the distribution of $\widehat{p}_{kl,kl}(7)$]

The distributions of the elements of $\widetilde{p}_{kl,kl}(7)$ are more skewed to the right, and assign a larger probability mass to values close to zero.

Let us now consider the dispersion of the estimators. The distribution of $\widehat{p}_{kl,kl}(7)$ is more concentrated around the mean value, as deduced by comparing the standard deviations and the interquantile ranges, or the histograms of the estimators. It is important to note that the time averaged estimator, which is more precise than the cross-sectional one, is itself not very accurate.

Finally, we can compare the results for migration correlations at horizon $h = 7$ obtained using estimators $\widetilde{\rho}_{kl,kl}(h)$ and $\widehat{\rho}_{kl,kl}(h)$ [see equations (15) and (16)], which are displayed in Tables 11 and 12, respectively. The true values of migration correlations are reported in Table 13. The empirical distributions of the estimators $\widetilde{\rho}_{kl,kl}(h)$ and $\widehat{\rho}_{kl,kl}(h)$ are reported in Figures 6 and 7, respectively.

[Insert Figure 6: Histogram of the distribution of $\widetilde{\rho}_{kl,kl}(h)$]

[Insert Figure 7: Histogram of the distribution of $\widehat{\rho}_{kl,kl}(h)$]

The proposed estimator $\widehat{\rho}_{kl,kl}(h)$ taking into account the Markov property is both less biased and more accurate. In particular, migration correlations estimated using the standard estimator $\widetilde{\rho}_{kl,kl}(h)$ are severely downward biased. Their use for computing a Credit VaR will automatically imply significant underestimation of the required capital.

[17] The (finite sample) bias of estimator $\widehat{p}_{kl,kl}(7)$ is due to the fact that it is the power of an unbiased estimator: $\widehat{P}(7) = \widehat{P}^7$.

8 Concluding remarks

The aim of this paper was to define precisely the notion of migration correlation, in order to study the properties of the cross-sectional estimators and compare them with alternative estimators taking into account the time dimension and the dynamic properties of the underlying model. It has been explained why the cross-sectional estimator is not consistent and the Monte-Carlo study points out its poor performance even if the cross-sectional dimension is large. In fact these basic cross-sectional estimators can be used to deduced the efficient time averaged estimator. We have also explained how to construct simultaneously the estimators of migration correlations at different horizons without loosing the information contained in the time dimension and in the underlying model. Since the estimated migration correlations are important tools for evaluating the risk on a credit portfolio, the estimation bias or the lack of accuracy will induce errors in the corresponding required capital. In practice by underestimating the true migration correlation, the estimated required capital will also be too small.

Since mainly the time dimension matters (not the cross-sectional dimension), there is a need for rather long individual rating histories. In practice the data bases of Moody's, Standard & Poor's, the French Central Bank, etc include between 15 and 20 years of reliable data. This is a minimal number to estimate the one-year migration correlation under an i.i.d. assumption on the underlying stochastic transition matrix. This is not sufficient to test this assumption of flat term structure of migration correlation with a sufficient power or to introduce more complicated dynamics on these transition matrices. This lack of time dimension is ever increased when some macroeconomic effects on transitions are introduced, for instance when different distributions of transition matrices are considered for recession and expansion periods [see e.g. Nickell, Perraudin, Varotto (2000), Bangia et alii (2002), and the discussion in Klaassen, Lucas (2002)]. This approach is much more appropriate for retail credits, such as classical consumption credit, mortgages, revolving credit, ... for which internal scores and ratings can be available on a monthly basis, with a time dimension of $T = 150 - 200$.

REFERENCES

Albanese, C., Campolieti, J., Chen, O., and A., Zavidonov (2003): "Credit Barrier Models", University of Toronto DP.

Bangia, A., Diebold, F., Krominus, A., Schlagen, C., and T., Schuermann (2002): "Ratings Migration and Business Cycle, with Application to Credit Portfolio Stress Testing", Journal of Banking and Finance, 26, 445-474.

Bardos, M., Foulcher, S., and E., Bataille (2004): "Les Scores de la Banque de France: Methode, Resultats, Applications", Observatoire des Entreprises, Banque de France.

Basle Committee on Banking Supervision (2002): "Quantitative Impact Study 3. Technical Guidance", Bank for International Settlements, Basel.

Brady, B., and R., Bos (2002): "Record Defaults in 2001. The Results of Poor Credit Quality and a Weak Economy", Special Report, Standard & Poor's, February.

Brady, B., Vazza, D., and R., Bos (2003): "Ratings Performance 2002: Default, Transition, Recovery and Spreads", Website of Standard & Poor's.

Carty, L. (1997): "Moody's Rating Migration and Credit Quality Correlation, 1920-1996", Moody's Investor Service.

Crouhy, M., Galai, D., and R. Mark (2000): "A Comparative Analysis of Current Credit Risk Models", Journal of Banking and Finance, 29, 59-117.

Das, S., and G., Geng (2002): "Modelling the Processes of Correlated Default", Working Paper, Santa Clara University.

Davis, M. (1999): "Contagion Modelling of Collateralized Bond Obligations", Working Paper, Bank of Tokyo, Mitsubishi.

Davis, M., and V., Lo (2001): "Infectious Defaults", Quantitative Finance, 1, 382-387.

De Servigny, A., and O., Renault (2002): "Default Correlation: Empirical Evidence", Standard and Poor's, September.

Duffie, D., and K., Singleton (1999): "Modeling the Term Structure of Defaultable Bonds", Review of Financial Studies, 12, 687-720.

Erturk, E. (2000): "Default Correlation Among Investment Grades Borrowers", Journal of Fixed Income, 9, 55-60.

Feng, D., Gouriéroux, C., and J. Jasiak (2003): "The Ordered Qualitative Model for Rating Transitions", Discussion Paper, e-Finance Group, Montreal.

Foulcher, S., Gouriéroux, C., and A., Tiomo (2003): "La structure par terme des taux de défauts et de Ratings", Banque de France DP.

Frey, R., and A., McNeil (2001): "Modelling Dependent Defaults", Working Paper, ETH Zurich.

Gagliardini, P., and C., Gouriéroux (2003): "Stochastic Migration Models", CREST DP.

Gieseke, K., (2001): "Correlated Defaults with Incomplete Information", Working Paper, Humboldt University.

Gordy, M. (1998): "A Comparative Anatomy of Credit Risk Models", Journal of Banking and Finance, 24, 119-149.

Gordy, M., and E., Heitfield (2002): "Estimating Default Correlation from Short Panels of Credit Rating Performance Data", Working Paper, Federal Reserve Board.

Gouriéroux, C. (2000): *Econometrics of Qualitative Dependent Variables*, Cambridge University Press.

Gouriéroux, C., and A., Monfort (2002): "Equidependence in Qualitative and Duration Models with Application to Credit Risk", CREST DP.

Gouriéroux, C., Monfort, A., and V., Polimenis (2003): "Affine Models for Credit Risk", CREST DP.

Gupton, G., Finger, C., and M., Bhatia (1997): "Creditmetrics-technical document", Technical Report, the RiskMetrics Group.

Kish, L. (1967): *Survey Sampling*, Wiley.

Konijn, H. S. (1973): *Statistical Theory of Sample Survey Design and Analysis*, North Holland.

Klaassen, P., and A., Lucas (2002): "Dynamic Credit Risk Modelling", Discussion Paper.

Lando, D. (1998): "On Cox Processes and Credit Risky Securities", Review of Derivatives Research, 2, 99-120.

Li, D. (2000): "On Default Correlation: A Copula Approach", Journal of Fixed Income, 9, 43-54.

Lucas, A., Klaassen, P., Spreij, P., and S., Straetmans (2001): "An Analytic Approach to Credit Risk of Large Bond and Loan Portfolios", Journal of Banking and Finance, 25, 1635-1664.

Lucas, D. (1995): ”Default Correlation and Credit Analysis”, Journal of Fixed Income, March, 76-87.

Maddala, G. S. (1986): *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

Merton, R. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", Journal of Finance, 20, 449-470.

Nagpal, K., and R., Bahar (1999): "An Analytical Approach for Credit Risk Analysis Under Correlated Defaults", Credit Metrics Monitor, 51-79, April.

Nagpal, K., and R., Bahar (2000): "Credit Risk Modelling in Presence of Correlations, Part 2; An Analytical Approach", Standard & Poor's.

Nagpal, K., and R., Bahar (2001)a: "Credit Risk in Presence of Correlations, Part 1; Historical Data for US Corporates", Risk, 14, 129-132.

Nagpal, K., and R., Bahar (2001)b: "Measuring Default Correlation", Risk, 14, 129-132.

Nickell, P., Perraudin, W., and S., Varotto (2000): "Stability of Rating Transitions", Journal of Banking and Finance, 24, 203-227.

Yu, F. (2002): "Correlated Defaults in Reduced Form Models", Working Paper, University of California Irvine.

Wilson, T. (1997)a: "Portfolio Credit Risk, I", Risk, 10, 111-117.

Wilson, T. (1997)b: "Portfolio Credit Risk, II", Risk, 10, 56-61.

Zeng, B., and J., Zhang (2002): "Measuring Credit Correlations: Equity Correlations are not Enough!", Working Paper, KMV Corporation.

Zhou, C., (1997): "Default Correlation: An Analytical Result", Working Paper, Federal Reserve Board.

Zhou, C., (2001): "An Analysis of Default Correlation and Multiple Defaults", Review of Financial Studies, 19, 553-576.

Appendix 1
Markov properties

Let us consider the joint rating vector $Y_t = (Y_{1,t}, ..., Y_{n,t})$. Its transition is characterized by:

$$
\begin{aligned}
& P\left(Y_{1,t+1} = k_1^*, ..., Y_{n,t+1} = k_n^* \mid \underline{Y_{1,t}}, ..., \underline{Y_{n,t}}\right) \\
= \; & E\left[P\left(Y_{1,t+1} = k_1^*, ..., Y_{n,t+1} = k_n^* \mid \underline{Y_{1,t}}, ..., \underline{Y_{n,t}}, (\Pi_t)\right) \mid \underline{Y_{1,t}}, ..., \underline{Y_{n,t}}\right] \\
= \; & E\left[\pi_{k_1 k_1^*, t+1} ... \pi_{k_n k_n^*, t+1}\right], \quad \text{where } Y_{1,t} = k_1, ..., Y_{n,t} = k_n.
\end{aligned}
$$

We deduce that the process (Y_t) is a Markov process. By a similar argument, the bivariate process $(Y_{i,t}, Y_{j,t})$, t varying, is also a Markov process, with transition probabilities:

$$P\left[Y_{i,t+1} = k^*, Y_{j,t+1} = l^* \mid Y_{i,t} = k, Y_{j,t} = l\right] = E\left[\pi_{kk^*,t}\pi_{ll^*,t}\right].$$

Table 1

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	90.26	3.59	0.00	0.51	0.00	0.00	0.00	0.00
AA	0.16	89.42	5.69	0.32	0.00	0.00	0.00	0.16
A	0.00	2.11	86.59	5.28	0.08	0.08	0.00	0.08
BBB	0.00	0.37	3.75	85.82	3.48	0.00	0.00	0.18
BB	0.00	0.00	0.13	2.77	80.10	6.17	0.50	1.13
B	0.00	0.00	0.22	0.33	2.21	77.65	3.65	6.97
CCC	0.00	0.00	0.00	2.70	0.00	2.70	55.41	31.08
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table 1: One-year transition probabilities (in %) for year 2001 drawn from the rating histories of 9769 obligors rated by Standard & Poor's [from Brady, Bos (2002)]. The transition probabilities on a row fail to sum to 100.00, since some obligors migrated in the non-rated category at the end of the year.

Table 2

Mean	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.633	0.112	0.000
$(2,2)$	0.114	0.505	0.024
$(3,3)$	0	0	1

Median	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.635	0.108	0.000
$(2,2)$	0.110	0.507	0.020
$(3,3)$	0	0	1

Std. dev.	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.069	0.040	0.001
$(2,2)$	0.043	0.053	0.017
$(3,3)$	–	–	–

MSE	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.005	0.002	0.000
$(2,2)$	0.002	0.003	0.000
$(3,3)$	–	–	–

5% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.517/0.744	0.053/0.182	0.000/0.001
$(2,2)$	0.051/0.190	0.416/0.590	0.004/0.057
$(3,3)$	–	–	–

1% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.462/0.782	0.035/0.225	0.000/0.004
$(2,2)$	0.032/0.234	0.375/0.618	0.002/0.079
$(3,3)$	–	–	–

Table 2: Estimated joint transition probabilities using estimator $\widehat{p}_{kk^*,ll^*}$ defined in (10). In the first four tables we report the average, the median, the standard deviation and the Mean Squared Error, respectively, of the estimator $\widehat{p}_{kk^*,ll^*}$ in the Monte-Carlo sample. The ranges between the 0.05 and the 0.95, resp. 0.01 and 0.99, quantiles of the estimator are reported in the last two tables. The rows correspond to the starting grades (k,k), the columns to the final grades (k^*,k^*), $k,k^*=1,2,3$.

Table 3

Mean	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.634	0.109	0.000
$(2,2)$	0.115	0.504	0.023
$(3,3)$	0	0	1

Median	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.710	0.024	0.000
$(2,2)$	0.024	0.572	0.000
$(3,3)$	0	0	1

Std. dev.	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.306	0.173	0.006
$(2,2)$	0.189	0.234	0.074
$(3,3)$	–	–	–

MSE	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.093	0.030	0.000
$(2,2)$	0.035	0.055	0.005
$(3,3)$	–	–	–

5% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.064/0.997	0/0.534	0/0.000
$(2,2)$	0/0.558	0.056/0.772	0/0.131
$(3,3)$	–	–	–

1% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.007/1	0/0.726	0/0.002
$(2,2)$	0/0.833	0.006/0.823	0/0.391
$(3,3)$	–	–	–

Table 3: Estimated joint transition probabilities using estimator $\widehat{p}_{kl,t}$ defined in (1) at date $t = 10$. In the first four tables we report the average, the median, the standard deviation and the Mean Squared Error, respectively, of the estimator $\widehat{p}_{kl,t}$ in the Monte-Carlo sample. The range between the 0.05 and the 0.95, resp. 0.01 and 0.99, quantiles of the estimator are reported in the last two tables. The rows correspond to the starting grades (k, k), the columns to the final grades (l, l), $k, l = 1, 2, 3$.

Table 4

	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.634	0.109	0.000
$(2,2)$	0.113	0.505	0.023
$(3,3)$	0	0	1

Table 4: Joint transition probabilities p_{kk^*,kk^*}. The rows corresponds to the starting grades (k,k), the columns to the final grades (k^*,k^*), $k,k^* = 1,2,3$.

Table 5

Mean	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.297	0.285	0.025
$(2,2)$	0.295	0.179	0.210
$(3,3)$	0	0	1

Median	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.294	0.284	0.011
$(2,2)$	0.292	0.176	0.191
$(3,3)$	0	0	1

Std. dev.	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.083	0.075	0.044
$(2,2)$	0.084	0.062	0.107
$(3,3)$	–	–	–

MSE	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.007	0.006	0.004
$(2,2)$	0.007	0.004	0.012
$(3,3)$	–	–	–

5% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.166/0.438	0.165/0.410	0.002/0.096
$(2,2)$	0.163/0.439	0.083/0.287	0.070/0.413
$(3,3)$	–	–	–

1% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.124/0.502	0.123/0.466	0.001/0.214
$(2,2)$	0.117/0.498	0.056/0.344	0.044/0.515
$(3,3)$	–	–	–

Table 5: Estimated migration correlations using estimator $\widehat{\rho}_{kk^*,ll^*}$ defined in (11). In the first four tables we report the average, the median, the standard deviation and the Mean Squared Error, respectively, of the estimator $\widehat{\rho}_{kk^*,ll^*}$ in the Monte-Carlo sample. The ranges between the 0.05 and the 0.95, resp. 0.01 and 0.99, quantiles of the estimator are reported in the last two tables. The rows correspond to the starting grades (k,k), the columns to the final grades (k^*,k^*), $k,k^*=1,2,3$.

Table 6

Mean	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	-0.007	-0.007	-0.007
$(2,2)$	-0.007	-0.007	-0.007
$(3,3)$	0	0	1

Median	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	-0.003	-0.003	-0.003
$(2,2)$	-0.003	-0.003	-0.003
$(3,3)$	0	0	1

Std. dev.	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.029	0.029	0.029
$(2,2)$	0.026	0.026	0.026
$(3,3)$	$-$	$-$	$-$

MSE	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.099	0.091	0.007
$(2,2)$	0.098	0.037	0.058
$(3,3)$	$-$	$-$	$-$

5% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	$-0.019/-0.001$	$-0.019/-0.001$	$-0.019/-0.001$
$(2,2)$	$-0.019/-0.002$	$-0.019/-0.002$	$-0.019/-0.002$
$(3,3)$	$-$	$-$	$-$

1% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	$-0.067/-0.001$	$-0.067/-0.001$	$-0.067/-0.001$
$(2,2)$	$-0.063/-0.001$	$-0.063/-0.001$	$-0.063/-0.001$
$(3,3)$	$-$	$-$	$-$

Table 6: Estimated migration correlations using estimator $\widehat{\rho}_{kl,t}$ defined in (2) at date $t = 10$. In the first four tables we report the average, the median, the standard deviation and the Mean Squared Error, respectively, of the estimator $\widehat{\rho}_{kl,t}$ in the Monte-Carlo sample. The range between the 0.05 and the 0.95, resp. 0.01 and 0.99, quantiles of the estimator are reported in the last two tables. The rows correspond to the starting grades (k,k), the columns to the final grades (l,l), $k,l = 1,2,3$.

Table 7

	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.305	0.293	0.072
$(2,2)$	0.305	0.184	0.232
$(3,3)$	$-$	$-$	$-$

Table 7: Migration correlations ρ_{kk^*,kk^*}. The rows corresponds to the starting grades (k,k), the columns to the final grades (k^*,k^*), $k,k^*=1,2,3$.

Table 8

Mean	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.268	0.174	0.058
$(2,2)$	0.201	0.136	0.137
$(3,3)$	0	0	1

Median	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.255	0.176	0.035
$(2,2)$	0.182	0.137	0.103
$(3,3)$	0	0	1

Std. dev.	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.129	0.048	0.065
$(2,2)$	0.118	0.039	0.102
$(3,3)$	–	–	–

MSE	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.017	0.002	0.004
$(2,2)$	0.014	0.002	0.010
$(3,3)$	–	–	–

5% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.081/0.499	0.091/0.251	0.003/0.193
$(2,2)$	0.043/0.423	0.070/0.200	0.017/0.340
$(3,3)$	–	–	–

1% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.040/0.606	0.060/0.279	0.001/0.300
$(2,2)$	0.019/0.540	0.046/0.223	0.006/0.454
$(3,3)$	–	–	–

Table 8: Estimated joint transition probabilities at horizon $h = 7$ using estimator $\widetilde{p}_{kk^*,ll^*}(7)$ defined in (13). In the first four tables we report the average, the median, the standard deviation and the Mean Squared Error, respectively, of the estimator $\widetilde{p}_{kk^*,ll^*}(7)$ in the Monte-Carlo sample. The ranges between the 0.05 and the 0.95, resp. 0.01 and 0.99, quantiles of the estimator are reported in the last two tables. In each table the rows correspond to the starting grades (k,k), the columns to the final grades (k^*,k^*), $k,k^* = 1,2,3$.

Table 9

Mean	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.277	0.168	0.066
$(2,2)$	0.213	0.131	0.137
$(3,3)$	0	0	1

Median	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.268	0.169	0.051
$(2,2)$	0.201	0.131	0.120
$(3,3)$	0	0	1

Std. dev.	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.110	0.031	0.057
$(2,2)$	0.102	0.023	0.089
$(3,3)$	–	–	–

MSE	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.012	0.001	0.003
$(2,2)$	0.011	0.001	0.008
$(3,3)$	–	–	–

5% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.114/0.476	0.115/0.218	0.007/0.181
$(2,2)$	0.072/0.404	0.095/0.170	0.025/0.311
$(3,3)$	–	–	–

1% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.072/0.569	0.086/0.239	0.002/0.256
$(2,2)$	0.043/0.502	0.078/0.189	0.011/0.404
$(3,3)$	–	–	–

Table 9: Estimated joint transition probabilities at horizon $h = 7$ using estimator $\widehat{p}_{kk^*,ll^*}(7)$ defined in (14). In the first four tables we report the average, the median, the standard deviation and the Mean Squared Error, respectively, of the estimator $\widehat{p}_{kk^*,ll^*}(7)$ in the Monte-Carlo sample. The ranges between the 0.05 and the 0.95, resp. 0.01 and 0.99, quantiles of the estimator are reported in the last two tables. In each table the rows correspond to the starting grades (k, k), the columns to the final grades (k^*, k^*), $k, k^* = 1, 2, 3$.

Table 10

	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.265	0.173	0.058
$(2,2)$	0.198	0.135	0.131
$(3,3)$	0	0	1

Table 10: Joint transition probabilities $p_{kk^*,ll^*}(7)$ at horizon $h = 7$. The rows corresponds to the starting grades (k, k), the columns to the final grades (k^*, k^*), $k, k^* = 1, 2, 3$.

Table 11

Mean	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.200	0.134	0.096
$(2,2)$	0.172	0.128	0.126
$(3,3)$	0	0	1

Median	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.186	0.127	0.068
$(2,2)$	0.157	0.121	0.103
$(3,3)$	0	0	1

Std. dev.	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.089	0.059	0.089
$(2,2)$	0.083	0.058	0.092
$(3,3)$	–	–	–

MSE	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.011	0.004	0.015
$(2,2)$	0.011	0.003	0.014
$(3,3)$	–	–	–

5% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.081/0.365	0.049/0.241	0.012/0.277
$(2,2)$	0.064/0.331	0.049/0.236	0.027/0.307
$(3,3)$	–	–	–

1% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.053/0.465	0.030/0.297	0.006/0.421
$(2,2)$	0.042/0.422	0.030/0.294	0.015/0.445
$(3,3)$			

Table 11: Estimated migration correlations at horizon $h = 7$ using estimator $\widetilde{\rho}_{kk^*,ll^*}(7)$ defined in (15). In the first four tables we report the average, the median, the standard deviation and the Mean Squared Error, respectively, of the estimator $\widetilde{\rho}_{kk^*,ll^*}(7)$ in the Monte-Carlo sample. The ranges between the 0.05 and the 0.95, resp. 0.01 and 0.99, quantiles of the estimator are reported in the last two tables. In each table the rows correspond to the starting grades (k, k), the columns to the final grades (k^*, k^*), $k, k^* = 1, 2, 3$.

Table 12

Mean	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.243	0.143	0.159
$(2,2)$	0.222	0.135	0.179
$(3,3)$	0	0	1

Median	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.243	0.142	0.149
$(2,2)$	0.220	0.134	0.168
$(3,3)$	0	0	1

Std. dev.	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.053	0.040	0.074
$(2,2)$	0.051	0.037	0.077
$(3,3)$	–	–	–

MSE	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.003	0.002	0.006
$(2,2)$	0.003	0.001	0.007
$(3,3)$	–	–	–

5% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.155/0.331	0.079/0.211	0.053/0.293
$(2,2)$	0.142/0.309	0.077/0.198	0.071/0.320
$(3,3)$	–	–	–

1% Interquantile	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.121/0.368	0.057/0.242	0.030/0.356
$(2,2)$	0.111/0.351	0.056/0.230	0.045/0.393
$(3,3)$	–	–	–

Table 12: Estimated migration correlations at horizon $h = 7$ using estimator $\widehat{\rho}_{kk^*,ll^*}(7)$ defined in (16). In the first four tables we report the average, the median, the standard deviation and the Mean Squared Error, respectively, of the estimator $\widehat{\rho}_{kk^*,ll^*}(7)$ in the Monte-Carlo sample. The ranges between the 0.05 and the 0.95, resp. 0.01 and 0.99, quantiles of the estimator are reported in the last two tables. In each table the rows correspond to the starting grades (k,k), the columns to the final grades (k^*,k^*), $k,k^* = 1,2,3$.

Table 13

	$(1,1)$	$(2,2)$	$(3,3)$
$(1,1)$	0.257	0.145	0.182
$(2,2)$	0.236	0.136	0.203
$(3,3)$	0	0	1

Table 13: Migration correlations $\rho_{kk^*,ll^*}(7)$ at horizon $h = 7$. The rows corresponds to the starting grades (k,k), the columns to the final grades (k^*,k^*), $k,k^* = 1,2,3$.

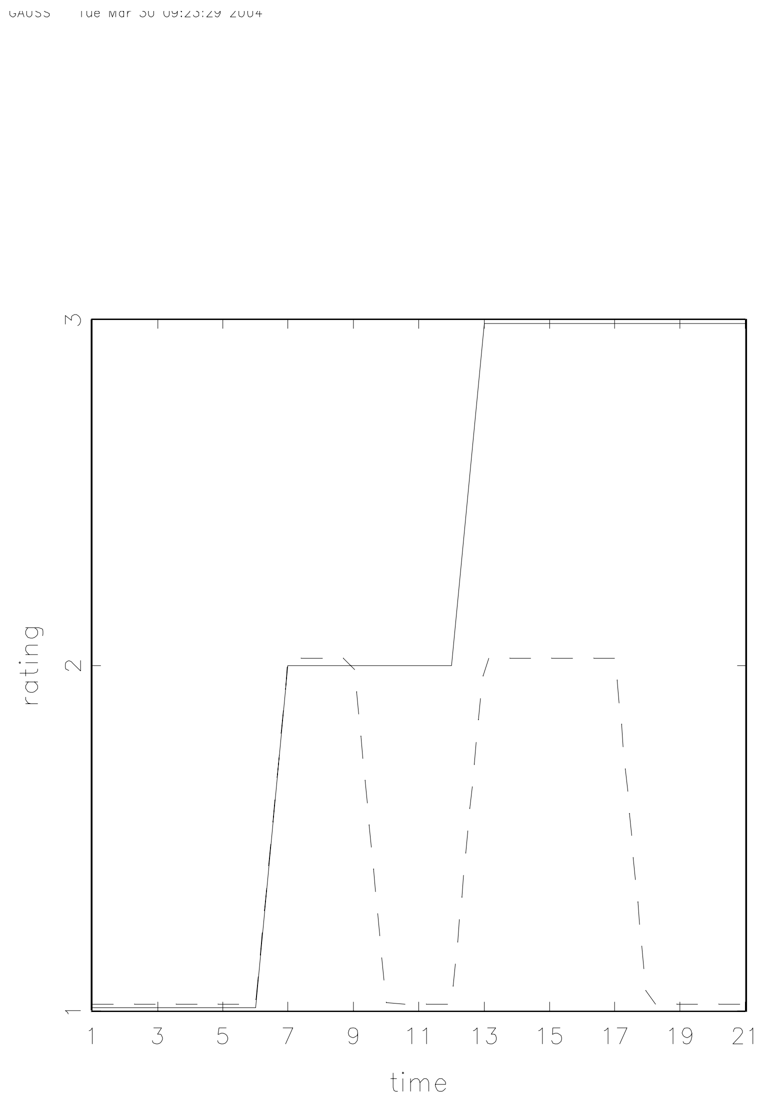

Figure 1: Rating histories of two firms.

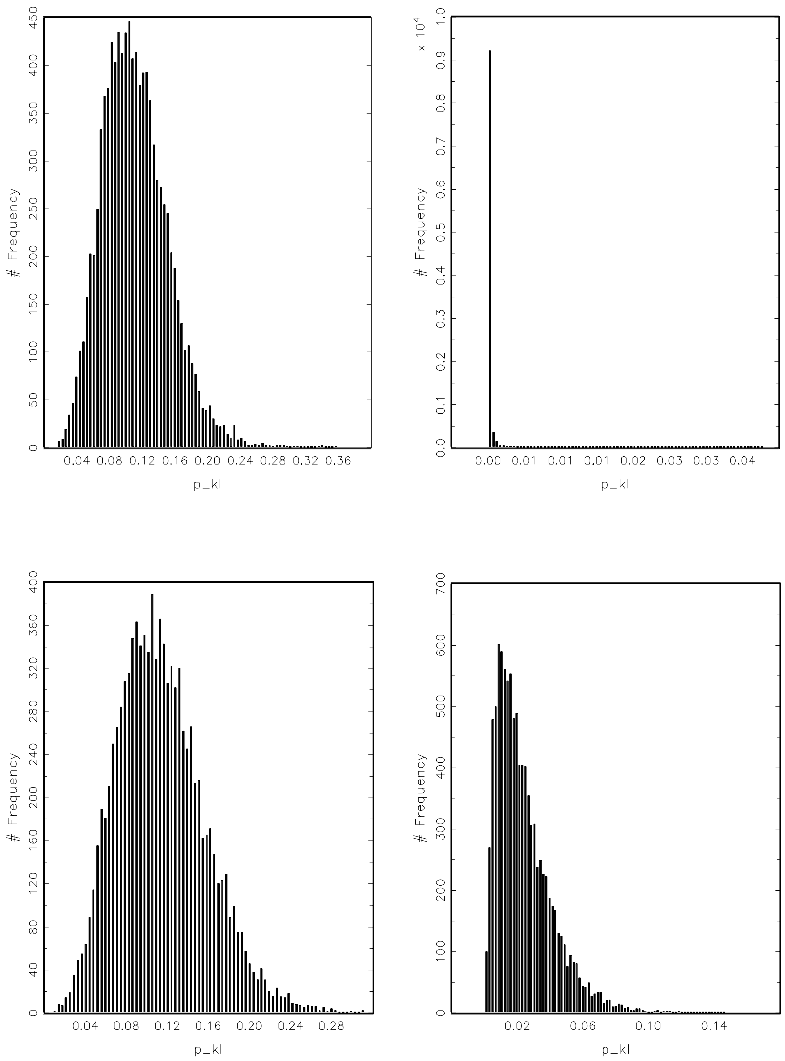

Figure 2: Histogram of $S = 10000$ Monte-Carlo replications of the estimator $\widehat{p}_{kk^*,ll^*}$ in (10). From left to right and top to bottom, the four panels report the estimators for $k = l = 1, k^* = l^* = 2$ (joint downgrade), $k = l = 1, k^* = l^* = 3$ (joint default starting from the best rating) $k = l = 2, k^* = l^* = 1$ (joint upgrade), and $k = l = 2, k^* = l^* = 3$ (joint default starting from the worst rating).

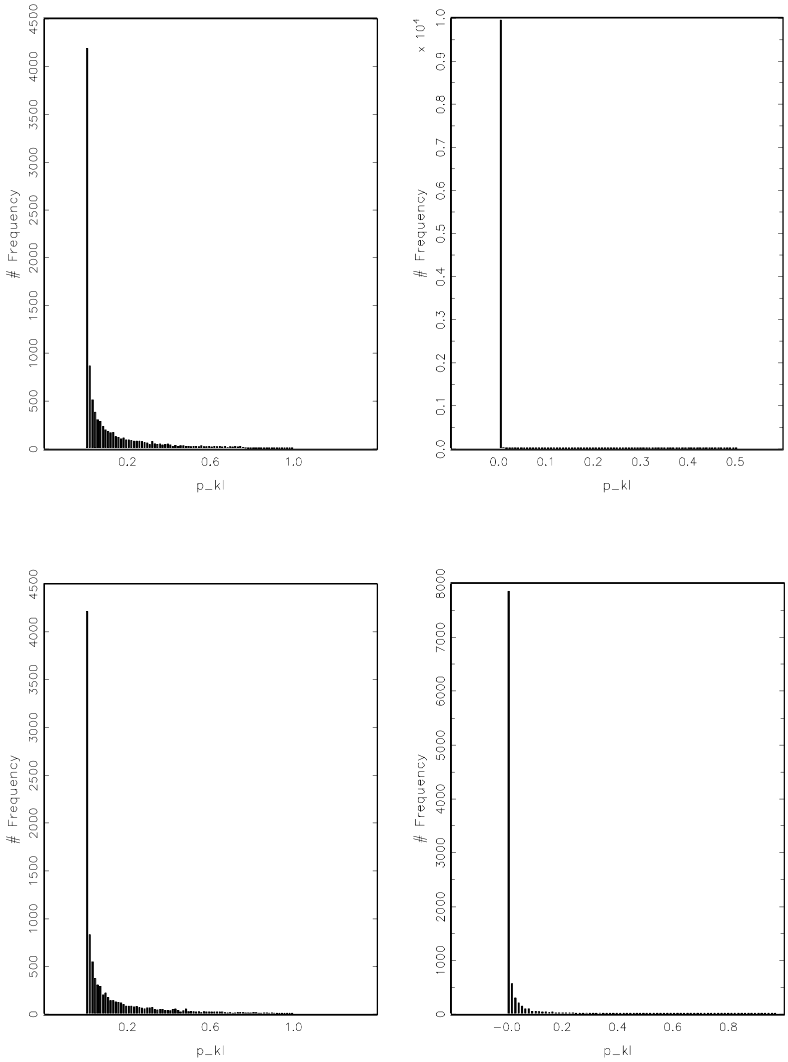

Figure 3: Histogram of $S = 10000$ Monte-Carlo replications of the estimator $\widehat{p}_{kl,t}$ in (1) for $t = 10$. From left to right and top to bottom, the four panels report the estimators for $k = 1, l = 2$ (joint downgrade), $k = 1, l = 3$ (joint default starting from the best rating) $k = 2, l = 1$ (joint upgrade), and $k = 2, l = 3$ (joint default starting from the worst rating).

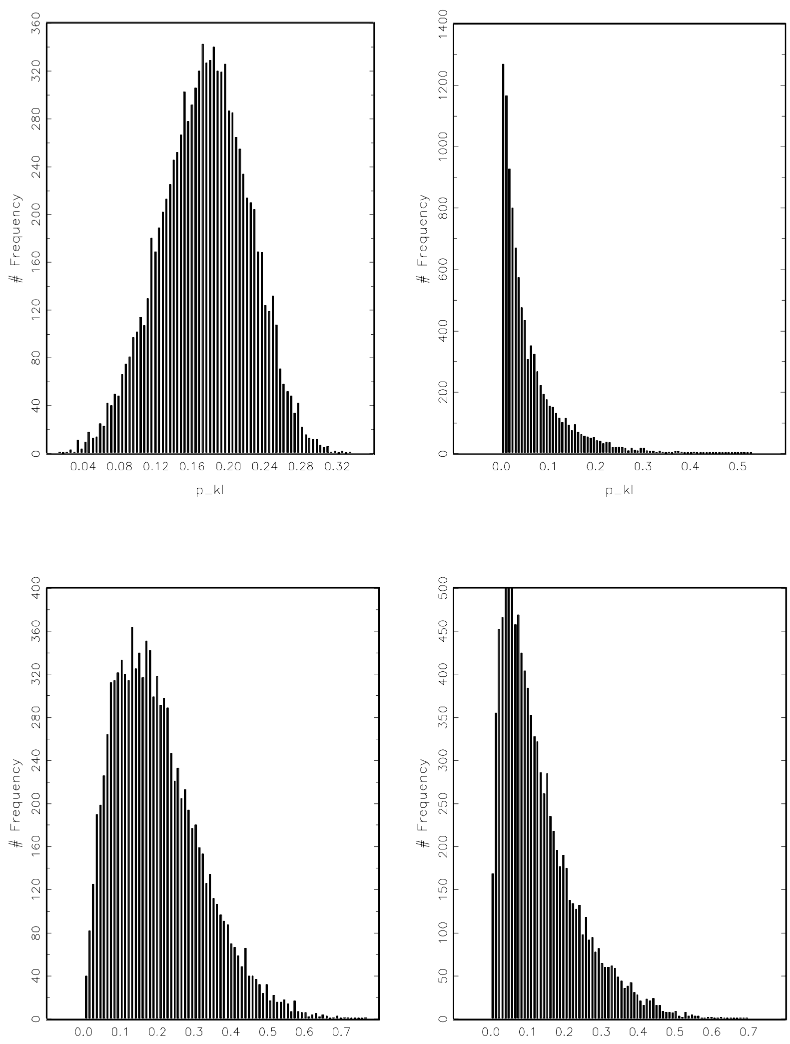

Figure 4: Histogram of $S = 10000$ Monte-Carlo replications of the estimator $\widetilde{p}_{kl,kl}(7)$ in (13). From left to right and top to bottom, the four panels report the estimators for $k = 1, l = 2$ (joint downgrade), $k = 1, l = 3$ (joint default starting from the best rating) $k = 2, l = 1$ (joint upgrade), and $k = 2, l = 3$ (joint default starting from the worst rating).

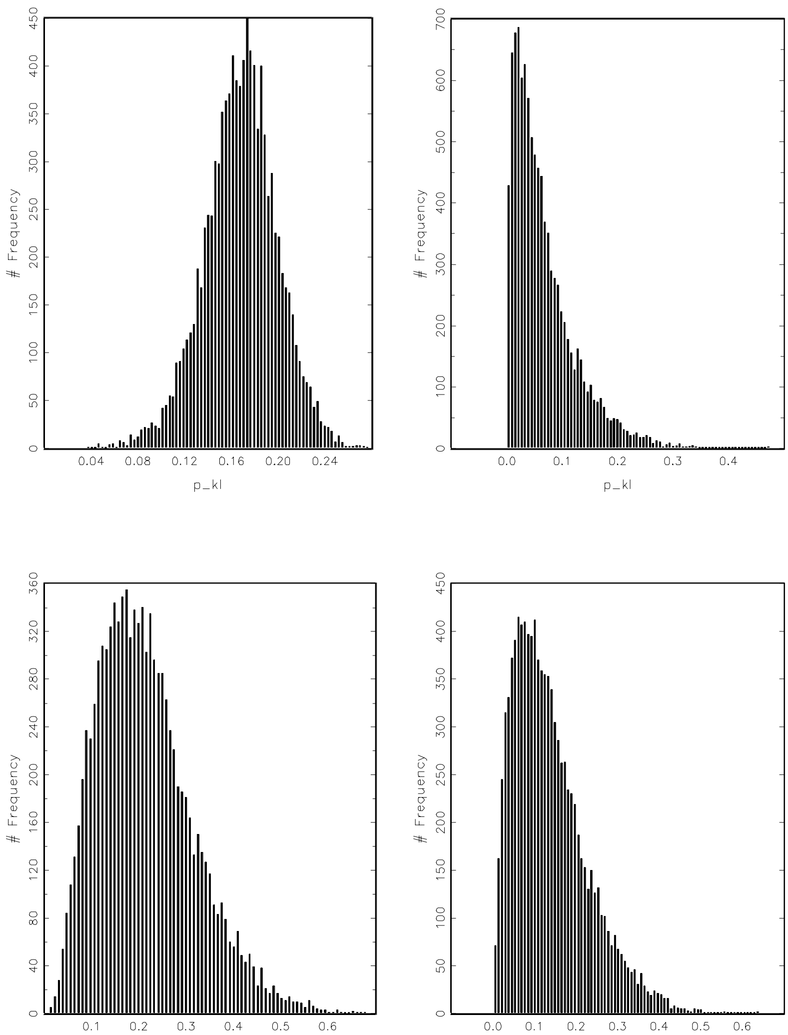

Figure 5: Histogram of $S = 10000$ Monte-Carlo replications of the estimator $\widehat{p}_{kl,kl}(7)$ in (14). From left to right and top to bottom, the four panels report the estimators for $k = 1, l = 2$ (joint downgrade), $k = 1, l = 3$ (joint default starting from the best rating) $k = 2, l = 1$ (joint upgrade), and $k = 2, l = 3$ (joint default starting from the worst rating).

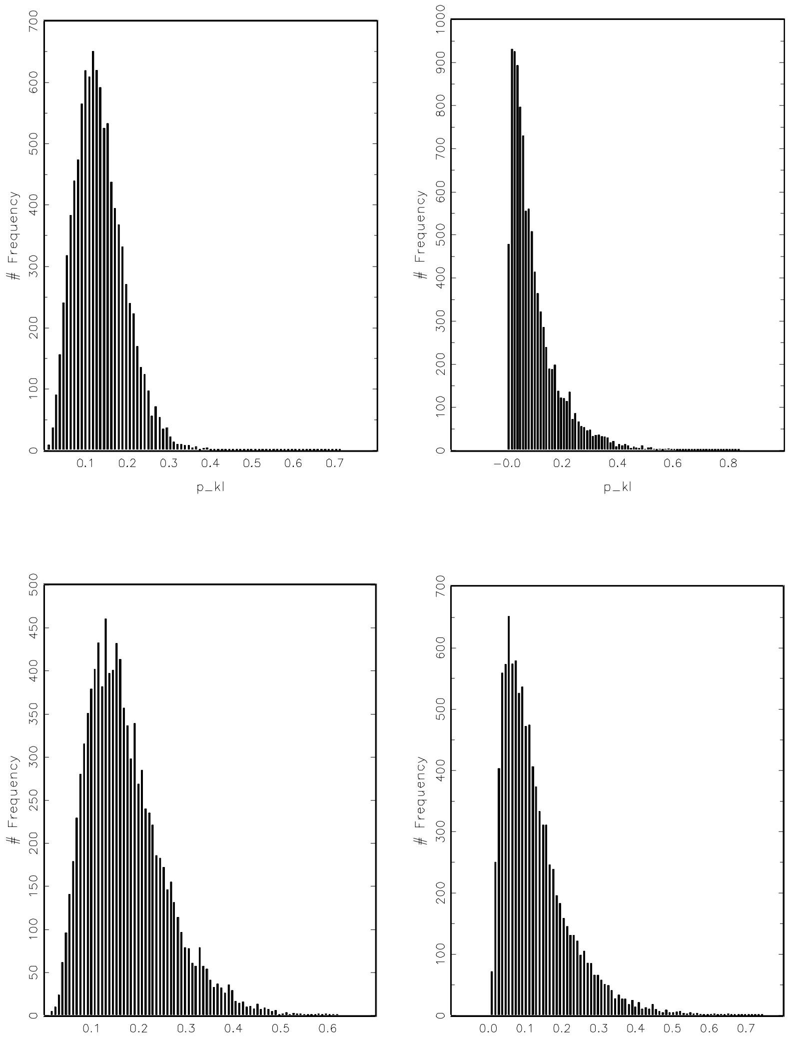

Figure 6: Histogram of $S = 10000$ Monte-Carlo replications of the estimator $\widetilde{\rho}_{kl,kl}(7)$ in (15). From left to right and top to bottom, the four panels report the estimators for $k = 1, l = 2$ (joint downgrade), $k = 1, l = 3$ (joint default starting from the best rating) $k = 2, l = 1$ (joint upgrade), and $k = 2, l = 3$ (joint default starting from the worst rating).

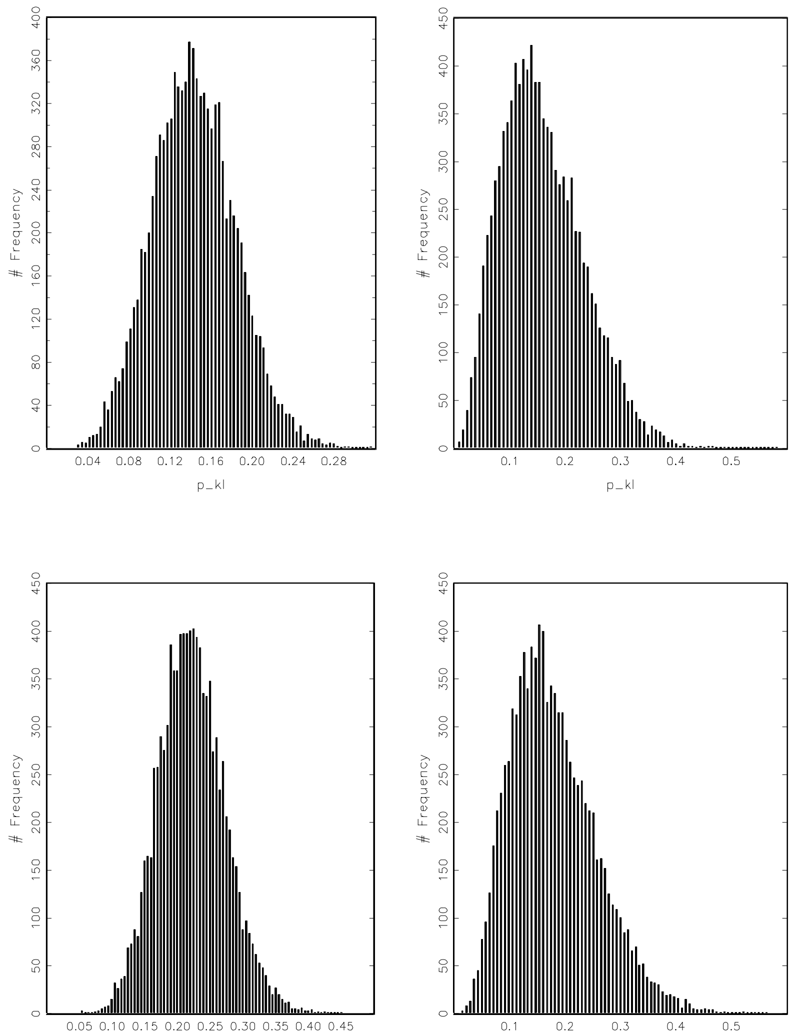

Figure 7: Histogram of $S = 10000$ Monte-Carlo replications of the estimator $\widehat{\rho}_{kl,kl}(7)$ in (16). From left to right and top to bottom, the four panels report the estimators for $k = 1, l = 2$ (joint downgrade), $k = 1, l = 3$ (joint default starting from the best rating) $k = 2, l = 1$ (joint upgrade), and $k = 2, l = 3$ (joint default starting from the worst rating).